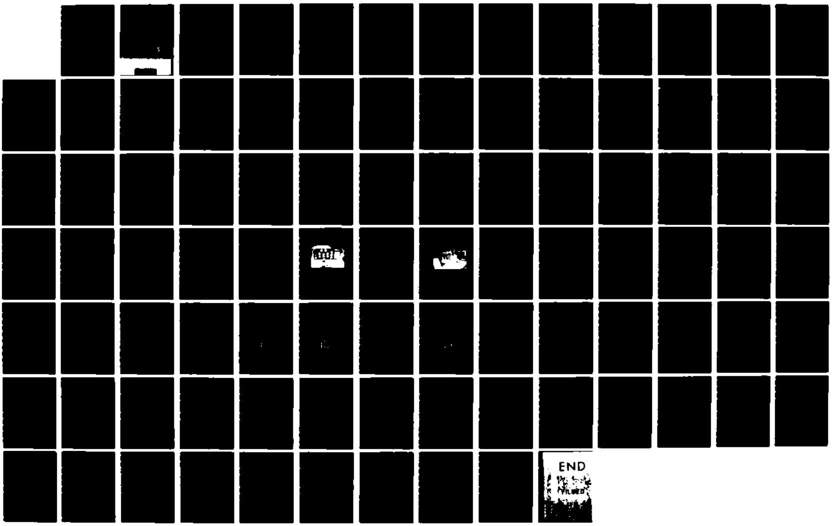
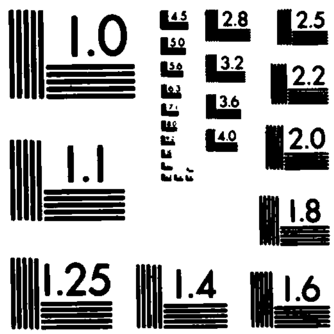
AD-A133 794 THE ABILITY TO PROCESS ABSTRACT INFORMATION(U) AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH M L MOROZE SEP 83 AFIT/CI/NR-83-51T 1/1

UNCLASSIFIED F/G 5/8 NL




1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

AD-A133 794

(1)

THE ABILITY TO PROCESS ABSTRACT INFORMATION

A Thesis Presented

By

Michael L. Moroze

Submitted to the Graduate School of the

University of Massachusetts in partial fulfillment

of the requirements for the degree of

MASTER OF SCIENCE

September 1983

Department of Industrial Engineering

and Operations Research

DEPARTMENT OF INDUSTRIAL ENGINEERING
AND OPERATIONS RESEARCH
UNIVERSITY OF MASSACHUSETTS AT AMHERST

UNCLASS

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFIT/CI/NR 83-51T	2. GOVT ACCESSION NO. AD-A133714	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) The Ability to Process Abstract Information		5. TYPE OF REPORT & PERIOD COVERED THESIS/~~DISSERTATION~~
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Michael L. Moroze		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS AFIT STUDENT AT: University of Massachusetts		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS AFIT/NR WPAFB OH 45433		12. REPORT DATE Sept 1983
		13. NUMBER OF PAGES 83
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASS
		15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

APPROVED FOR PUBLIC RELEASE: IAW AFR 190-17

Lynn Wolaver

LYNN E. WOLAVER
Dean for Research and
Professional Development

5 Oct 1983

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

ATTACHED

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASS

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

83 10 12 188

ABSTRACT

The increasing use of technology in human-machine systems has brought about the need to determine how the introduction of sophisticated technology impacts the human's performance. How this technology interacts with the stress, workload, and information processing capacity of the individual is discussed. The technology discussed is based on the use of advanced flight displays, particularly the Head-Up Display, in the aircraft environment. A study was conducted using three different methods of displaying the flight information to the operator. It was found that, although all subjects could perform their flight tasks within pre-determined criteria, under a loading condition performance using a relatively abstract presentation style was significantly poorer than performance using a more traditional, relatively concrete presentation style.

A

THE ABILITY TO PROCESS ABSTRACT INFORMATION

A Thesis Presented

By

Michael L. Moroze

Submitted to the Graduate School of the

University of Massachusetts in partial fulfillment

of the requirements for the degree of

MASTER OF SCIENCE

September 1983

Department of Industrial Engineering

and Operations Research

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

ACKNOWLEDGEMENTS

I would like to thank Capt. Bill Marshak at the United States Air Force Academy for his time and energy in developing and revising the data acquisition program used in this study and, also, I would like to thank Mark Normand for his perserverance and unending patience while developing the graphic displays used.

Many thanks also go out to Dr. Jefferson M. Koonce whose help and guidance throughout this project were invaluable.

INTRODUCTION

The performance of the human operator in man-machine systems has often been overlooked. Traditionally, the capabilities of the machine have been of primary concern. In the not so distant past this was a realistic viewpoint, the human's limitations were not as restrictive as the machine's. It was only logical, therefore, to place primary emphasis on the capabilities and limitations of the machine. Today, however, this traditional viewpoint has become obsolete. With today's technology the ability to produce machines that can overwhelm their operators has never been greater. It has become increasingly important to consider the human's limitations and capabilities in the design of man-machine systems.

The ability to present information about a particular system to the human operator has undergone a great deal of change in recent years. This change has resulted in the need to assess the information required by the operator to perform a task adequately. Several fundamental questions concerning information and the human operator include: What information does the operator require and what information should be directed to more automated sources? When should the operator receive the information? What form should the information take? These general questions are taking on increasing importance as we move into the information age.

If it has been decided that the operator is to receive the information, it is of critical importance to determine how to display the information. No matter when or how much information is displayed

to the operator, if it is not displayed in an understandable way, the operator will not be able to use it. Several factors influence the ability of the human operator to process information. These factors include the stress the operator feels, the inherent information processing ability of the operator, the workload of the operator, and the style of information presented to the operator. Although it is possible to think of these factors as separate entities, it is more realistic to view them as highly interrelated. It is for the sake of clarity and ease of understanding that they are presented separately.

This paper will discuss these fundamental factors and the impact they have on the man-machine sysem. The first part of this paper presents a review of stress, information processing and workload. After the introduction of these concepts, a discussion of their interaction with an individual's ability to handle information in one particular complex environment is presented. The complex environment chosen is modelled after that found in the pilot-aircraft environment. Finally a study conducted to assess how these factors interact with the human's ability to perform with different types of information is presented.

CHAPTER I

REVIEW OF THE LITERATURE

Stress

Like most of the topics discussed in this paper, stress is an ambiguous term. Although there has been a great deal of research into the area of stress, no unified theory or definition seems to exist. Stress can be thought of as a discipline in and of itself and could probably fill several volumes in its elaboration. Stress may take on several meanings depending on the population using the term. For example, stress to a physician takes on different connotations than stress to a psychologist or engineer. Even within the same audience stress will be viewed differently: to a clinical psychologist stress may be viewed in relation to personality variables whereas to an experimental psychologist stress may be viewed as a limit on performance, etc. It is easily seen, therefore, that stress is an extremely broad concept with many ramifications. Hogan and Hogan (in Alluisi and Fleishman, 1982) have introduced an admittedly arbitrary term to encompass this multifaceted concept. The Stress Activation Syndrome (SAS) has been offered as an adequate reference point which is meant to encompass all the connotations that the term stress has commonly included. According to Hogan and Hogan, SAS includes three components: (1) stressors, (2) psychological or subjective factors, and (3) the stress response.

Stressor

The stressor is considered the agent that may produce a stress response. A stressor may take on one (or both) of two forms, physical or psychological. A physical stressor is one that is found in the environment and acts upon the individual. Such stressors as extremes of temperature, loud noises, polluted air, or being struck by an object, (i.e., stick, bullet, truck, etc.) would be considered physical stressors (Hogan and Hogan). Although a physical stressor may occur at any time, Hogan and Hogan feel it is the psychological stressor that occurs on a day-in, day-out basis. A psychological stressor occurs with the anticipation of harm which may occur in two ways: (1) anticipation of physical harm, as in driving with a student driver, or (2) anticipation of social censure, as in fear of failure. These sources of stress may have a disruptive effect on the physiological and/or psychological processes of the individual. A stressful situation occurs whenever the normal relationship between an individual and his/her environment has been disrupted (Schaffer, 1954). However, it is the psychological perception or subjective evaluation by the individual that determines the effect the stressor will produce.

Perception of stress

Although the etiology of a stressful state may take on many forms, a critical element in labelling a situation as stressful resides within the perceptions of the individual. In order for the individual to perceive a particular situation as stressful, there must

be some perception of threat (Appley and Trumball, 1971; Hogan and Hogan, 1982; Lazarus, 1964; Lazarus, Deese, and Osler, 1952). The perception of threat is based upon some cognitive appraisal in which the situation before the individual is judged to be threatening for that individual. Therefore a stressful situation for one individual should not automatically be considered a stressful situation for another. As Appley and Trumball (1971, p. 592) state, "not only must a situation be of a given intensity to lead to stress, it must also be of a given kind for a particular person." As mentioned previously, the threat may be of two types: physical or psychological. Neither type of threat is more stressful than the other, rather it is the perception of the threat that enables us to judge one situation more threatening. A psychological stressor can be judged as powerful or devastating as a physical stressor (Thompson, 1975).

Stressors may have a disrupting effect on the physiological and/or psychological processes of the individual and the particular stress response depends on the individual. Although the physiological and psychological responses to a stressor are interrelated, they will be discussed separately below.

Physiological responses

Much of the research on stress is concerned with the physiological responses that are made in reaction to it. One of the first models that was offered was the General Adaptation Syndrome (GAS). Selye (1974) points out that this model reflects the stereotypical responses that are made by an organism when a stressor

is presented. These responses follow three stages. The alarm reaction which defines the first stage is a general mobilization of the body's resources in order to meet and handle the stressor(s) presented. Following the alarm reaction stage is the resistance stage. The resistance stage occurs when the stressor is prolonged and is characterized by the utilization of many vital substances in the body. When the production of these substances is unable to keep pace with their expenditure, the exhaustion stage is entered and finally the substances are depleted. Some of the effects of the GAS include marked changes in the nervous and endocrine system due to the imbalance of these different biochemicals. These changes include the secretion of or an increase in the secretion of norepinephrine, epinephrine, corticoids, growth hormones and various other biochemical substances (Schuler, 1980; Mulder, 1979; Ursin, Baade, and Levine, 1978).

The orientation reaction and defensive reaction are similar to the GAS. The reflexive orientation reaction causes an immediate response to any change in the environment and is further accompanied by several physiological changes such as desynchronization of the brain waves, a decrease in basal skin resistance, heart rate deceleration, etc. (Mulder, 1979). The defensive reaction prepares the organism for a fight or flight response by increasing the release of biochemical substances, similar to those found in the GAS such as epinephrine and norepinephrine, blood glucose, etc. (Mulder, 1979). Although specific effects of the various chemicals are not fully understood, various psychosomatic illnesses have been directly or

indirectly linked to these biochemical reactions to stress (see Figure 1).

Schaffer (1954) has offered another physiological response to stress. This early proposal suggested that, under stress, an individual fixates on whatever response is dominant at the time. This response is not extinguished even when followed by negative reinforcement. We become locked-in to a certain response which may or may not be appropriate. This is due to a process termed relative functional decortication. According to this process, Schaffer suggests the subcortical centers of the brain dominate the higher-level cortical centers. This is based on neurophysiological research in which decorticate subjects can quickly acquire a conditioned reflex, but do not have the ability to adapt and adopt appropriate behaviors as a normal subject would. This type of response has been observed in individuals under high-stress loads.

Psychological responses

Another model, offered by Hamilton (1975 in Mulder, 1979), defines anxiety (stress) as cognitive data that has been stored in long-term memory. These cognitive data include an avoidance component due to their cost which may take the form of embarrassment, physical danger, fear of failure, etc. When a stimulus causes the data to be recalled, the data are channelled to somewhere in the information processing system. Once there, they compete for space with whatever relevant information is presently being processed. Depending on the form of the information processing occurring (which will be discussed

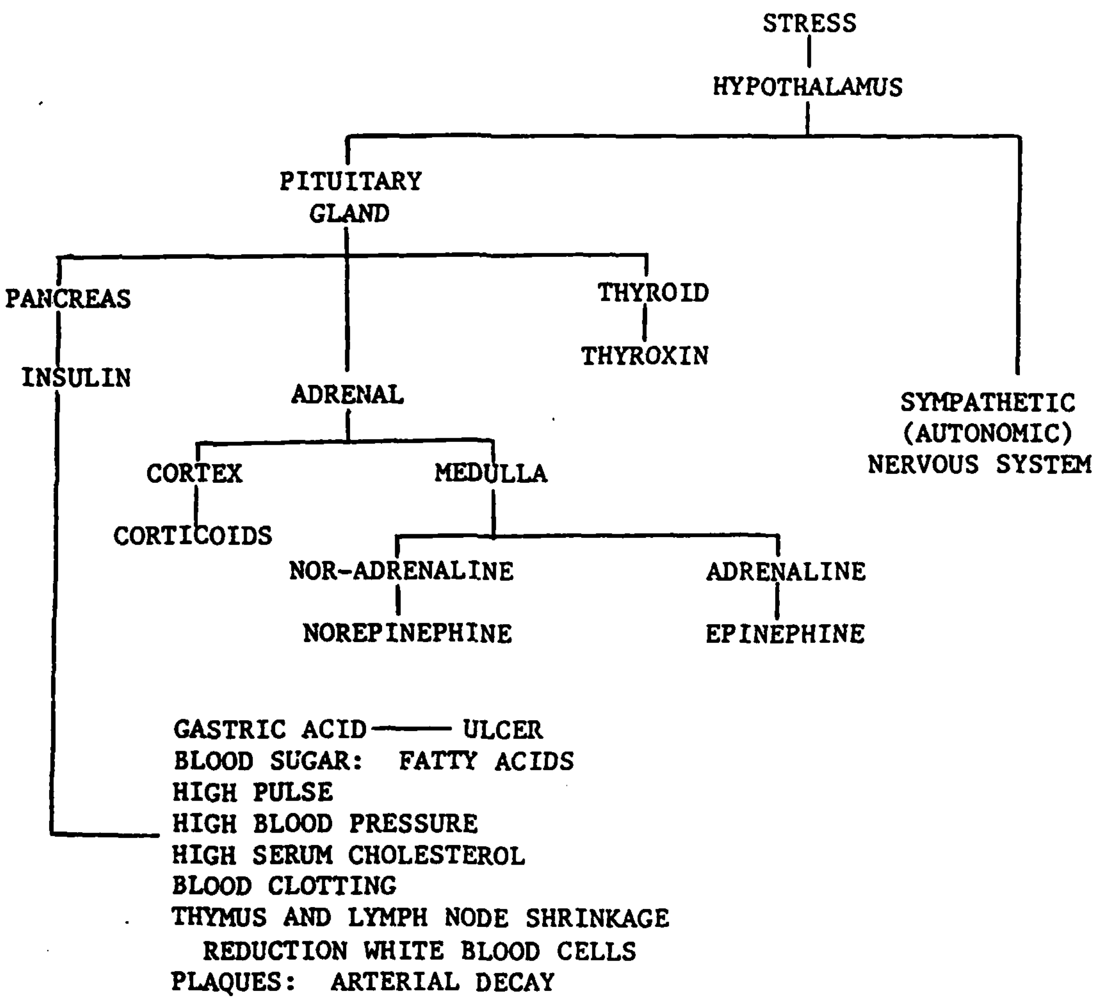


Figure 1. Biochemical and nervous responses associated with stress (Schuler, 1980, p. 202)

later), the anxiety-related cognitive data causes a decrease in the amount of information that can be processed. Also the negative feelings associated with the cost of the data may increase the overall feelings of stress felt by the individual. Therefore according to this model of stress and cognitive data, the cause of stress may be perceived in two ways: stress may be due either directly to the competition for processing space or indirectly to the automatic enabling of the negative feeling associated with the cost of the data.

The relationship of arousal (stress) to performance has been described as the now well-known inverted-U relationship. Low arousal is usually associated with poor performance and as arousal increases performance improves until an optimal point is reached, thereafter an increase in arousal causes performance to decrease. The poor performance associated with low arousal may be due to individuals omitting relevant data (i.e. not paying attention), whereas the poor performance associated with high arousal may be due to individuals including too much irrelevant data.

Another psychological response to high levels of stress has been offered by Broadbent (1971) and by Welford (1978). According to this model the individual experiencing high levels of stress does not consider all the relevant information available. Rather s/he will filter out information or selectively attend to the information which s/he subjectively determines to be most important. Thus peripheral information is not given importance and attention is shifted to the information felt to be most important. The performance decrement under high stress may be due to the information void incurred by not

attending to peripheral information which the individual may not think is important, but in actuality, it is. Peripheral information can also take the form of information presented in the periphery of a display.

Stress can have a greatly varied effect depending on the individual's perception. However once a situation is perceived as stressful, it is obvious that his or her performance will be affected whether it be in the short or long term. The amount of stress experienced and the interaction of stress with performance can have a great impact on an individual's behavior. Although stress has been shown to influence negatively both the physiological and psychological processes, Hamilton (1975, in Mulder, 1979) implies a direct cost to the information processing capability of the individual.

Information Processing

Information processing theory has borrowed from many fields in the course of its development. Such fields as communications engineering, information theory, linguistics, computer science, and engineering psychology have all made significant contributions to the development of information processing (Lachman, Lachman and Butterfield, 1979). The strength of the present-day information processing paradigm is based on the convergence of these areas and the broad scope they encompass. The relationship of these areas is easily recognized in the shared terminology. Such terms as channel capacity, buffer storage, processor, encode, etc. can all trace their beginnings

to one of these fields. Although the area of information processing has been widely researched, much of the research on information processing is based on a limited capacity model of the human operator. This limited capacity model can be traced back to Broadbent's model developed during the 1950's and later revised (1971). Another more recent model has been proposed by Schneider and Shiffrin (1977) and Shiffrin and Schneider (1977). Unlike Broadbent's filter models Schneider and Shiffrin propose two information processing modes: controlled and automatic processing.

Filter theories

Broadbent's theory proposes that information enters a buffer store which has an essentially unlimited capacity. Therefore the initial perceptual system is not subject to a selective process, rather the selectivity of the system follows the initial intake of information. Selective processes, likened to a filter, would allow some of the information to proceed through the system for further processing. The point here is that parallel processing occurs up to the point of the filter (i.e. the initial perceptual system) and, after the filter, processing only occurs on the information selected. This forms the basis for an all-or-none switch in which the filter allows for processing to take place only on the input "selected". This filter model is shown in Figure 2A.

According to this theory, any information on a nonattended channel will not be perceived by the individual. However, this does not account for the results that have suggested that information is

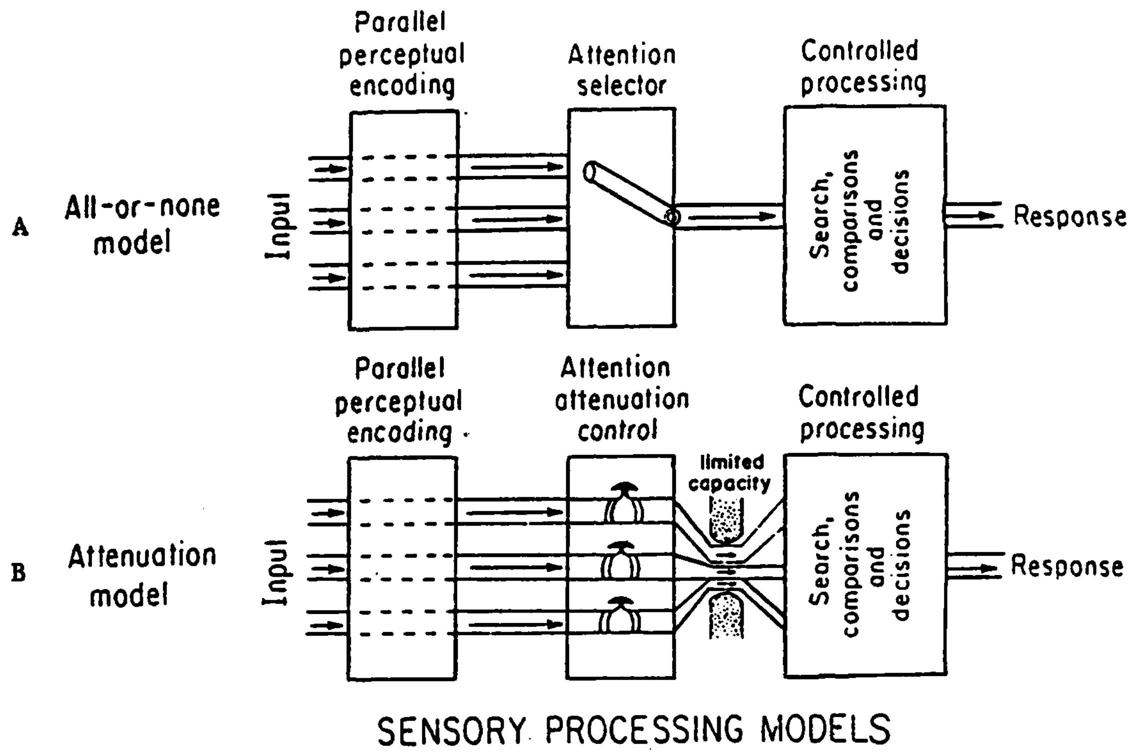


Figure 2. Filter Theories:
A. Broadbent's filter model
B. Treisman's attentuation model

(From Schneider and Shiffrin, 1977, p. 5)

perceived on nonattended channels. One such result is the so-called "cocktail party" effect (Hochberg, 1978). Specifically, when conversing with an individual at a party and attending to that individual, if your name is mentioned by someone else, you tend to perceive it. Although this type of response could be accounted for by the filter mechanism switching from an attended channel to a nonattended channel, a variation of this type of filter model seems to reflect the data better. Treisman (in Lachman, Lachman, and Butterfield, 1979) has proposed a model in which the filter has a limited capacity and is allocated by the subject to several input channels. This model has been labeled the attenuation model because the processing resources given to any particular channel are attenuated to the degree desired by the individual. This model is shown in Figure 2B.

Controlled and automatic processing

Although other theories and models exist, the two models discussed above form the foundation for most of the models presented in the literature. In a series of articles, Schneider and Shiffrin (1977; Shiffrin and Schneider, 1977) propose another information processing model. In many respects this model is in agreement with the previous models discussed, however some differences exist. According to this model, human information processing takes on two forms: controlled processing and automatic processing. It should be pointed out that other authors describe similar processes with various names, for example, the automatic and effortful processing described

by Hasher and Zacks (1979). However, the terms automatic and controlled processing will be used here due to their natural connotations.

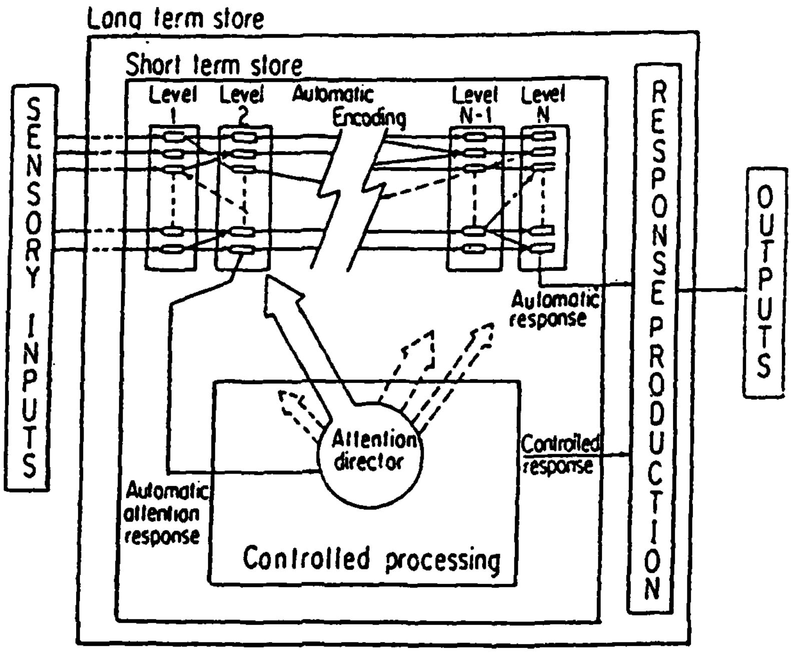
As the name implies, Schneider and Shiffrin suggest controlled processing occurs under the control and direction of the individual. Controlled processing is a temporary process which takes place in short-term store. Under the control and attention of the subject, a temporary activation of memory nodes takes place in a sequence that has not yet been learned. In this respect, controlled processing is relatively easy to set up, modify, and use in new situations. Controlled processing requires the attention and short-term capacity of the subject, and is often serial in nature. In visual and memory search tasks, the serial nature of the processing takes the form of a comparison process which takes place at a limited rate. The comparison process first compares a memory set item to all display items in turn and then chooses a new memory set item and continues until termination. Latency is a function not only of the duration of the comparison process, but also of the time it takes to choose a new memory set item.

Automatic processing is entirely different from controlled processing. Automatic processing does not take place under the subject's control, and it is learned or produced following the earlier use of controlled processing which has established a specific sequence for certain nodes. Once a sequence of nodes is learned in long-term store, the sequence can be triggered by the corresponding inputs and operates independently of the subject's control. Therefore, although

automatic processing may attract attention, it does not require the attention of the subject. This process requires a great deal of time and training to develop and, once developed, is difficult to suppress or modify. A model of this processing system has been illustrated by Shiffrin and Schneider (1977) and is shown in Figure 3.

Although Schneider and Shiffrin's model of the information processing system has garnered a great deal of support, some authors are less enthusiastic about the impact this model should have on current information processing theory development. While Schneider and Shiffrin suggest that their theory is a new formulation of information processing ability, Ryan (1983) puts forth a strong argument that it is merely a renaming of already well-established phenomena. Fisher (1982) argues that rather restrictive limits exist in situations where processing is believed to be automatic. Even considering these objections, as Ryan (1983) points out, it has been well-established "that human performance is load dependent in some cases and relatively load independent in others" (p. 171). Thus, at the very least, automatic and controlled processing can be used as synonyms for load independent and load dependent behavior, respectively.

Although it now seems that in all situations individuals may not be single-channel information processors and may not even be of limited capacity (Moray et al., 1979), use of the single-channel, limited-capacity model still appears to be appropriate in certain circumstances. An individual's ability to process information will have a great effect on the type, amount, and style of information



A model for automatic and controlled processing during tasks requiring detection of certain input stimuli. Short-term store is the activated subset of long-term store. N levels of automatic encoding are shown, the activated nodes being depicted within each level. The dashed arrows going from higher to lower levels indicate the possibility that higher level features can sometimes influence the automatic processing of lower level features. The solid arrow from a node in Level 2 to the attention system indicates that this node has produced an automatic-attention response, and the large arrow from the attention system to Level 2 indicates that the attention system has responded. The arrow from level N to the Response Production indicates that this node has called for an automatic overt response, which will shortly be executed. The arrow from Controlled Processing to the Response Production indicates the normal mode of responding in which the response is based on controlled comparisons and decisions. Were it not for the automatic responses indicated, detection would have proceeded in a serial, controlled search of nodes and levels in an order chosen by the subject.

Figure 3. Controlled and automatic processing (Shiffrin and Schneider, 1977, p. 163).

which can be comfortably presented. The ability to process and handle information is intricately related to the workload and feelings of workload experienced by an operator. The area of workload and, in particular, mental workload will be discussed in the next section.

Workload

Workload can be divided into two general cases: physical workload and mental workload. Although this separation is possible when discussing them, Moray et al. (1979) point out that there is no one observable task that is totally one or the other. All tasks include some part of each in the total contribution to the task, however the ratio of one to the other may vary greatly. Physical work is more easily observable and measurable and therefore it has become easier to define. Useful measurement tools have often revolved around the amount of oxygen consumed for a particular task. Each task requires a certain amount of oxygen consumed while attempting the task, and the maximum amount of oxygen that can be consumed is also limited to each individual (Mulder, 1979). By comparing the amount of oxygen consumed for a particular task to the maximum, a measure of physical workload is obtained.

Although we may have an intuitive feeling for what mental workload is, no single objective definition currently exists. Objective definitions exist only for specific tasks. While mental workload can be thought of as "how busy is the operator?" (Knowles, 1963), it is certainly a "multidimensional concept with many

definitions" (Meister, 1976, in Ogden, Levine, and Eisner, 1979, p. 529). These definitions can be applied only to mental workload in a general context, however, several factors must play a role in any definition of mental workload. These factors include the channel capacity, limited capacity processing, and effort employed by the operator (Sanders, 1979). Although the areas of channel capacity and limited capacity processing were previously discussed and may be of little consequence. under certain conditions and in certain situations, these models adequately describe the system under which individuals may work (such as in acquiring new skills, in novel situations, or in dynamic situations). Effort, on the other hand, was seen to be the amount of energy expended to accomplish the task, or *how hard the operator is trying.* As such, the amount of effort may be seen as the equivalent to the amount of mental workload. Just as there may be an individual limit on physical workload, effort may also be limited and be a function of the total amount of effort demanded at any particular moment in time.

Part of the problem in defining mental workload stems from the difficulty in measuring it. The measures employed are usually of two forms: behavioral measures and physiological measures. Willeges and Wierwille (1979) and Wierwille (1979), respectively, have reviewed the literature concerning behavioral and physiological measures. Their findings indicate that many of the measures presently used have not adequately proven their sensitivity to mental workload. Another literature review conducted by Ogden, Levine, and Eisner (1979) concentrated on the use of secondary tasks as a measure of mental

workload. They, also, found problems concerning the use of these methods in measuring workload due to changes in the strategy employed by the operator, to difficulties in eliminating structural limitations and central interference, and to limitations based on the single-channel, limited-capacity nature of the model. Physiological methods have an advantage of not relying on the single-channel, limited-capacity model, nor on interference (to some extent). However the equipment needed to gather the data and the difficulty in using the equipment are two inherent, realistic problems faced when using physiological measures. Also, many of the physiological measures measure the same attributes as are measured when researching stress so that the relationship between stress and mental workload is often confounded.

Human Performance Considerations

The previous sections give us a framework from which we can begin to analyze how such variables as stress, information processing, and workload affect human performance. Increases in technology have become increasingly more apparent in today's environment. It has been shown that the ability to perform a given task is dependent upon the amount of stress and workload felt by the operator and the operator's ability to process information obtained from the environment. Man has become increasingly dependent on displays to furnish information about the environment. With the advent of new technology and computers there has been a transition from the more traditional analog dial type

of displays to computer generated displays. The symbology used in the computer generated displays usually takes one of two forms: small pictures (pictographs) or abstract symbols. This transition has been seen in the design of international road signs, automobile instrument panels, and aircraft cockpits. Perhaps the greatest impact has been felt on the aircraft environment. The aircraft environment has seen an increasing sophistication, and as new technology has developed, the pilot has been increasingly taxed. With this increasing load on the pilot, there has come a need to aid the pilot in obtaining information, deciding on an action, and making the appropriate response. One such aid is the Head-Up Display (HUD). Although the HUD has been in use for a period of time, one area that may be a problem to its users is the symbology used. The *following sections* describe the HUD and some possible problems concerning its use.

General considerations of the head-up display

The HUD was developed during the 1950's to aid the pilot in obtaining and assimilating information. The major drive behind the HUD's development was the increasing performance capabilities and complexity of the newer aircraft. Due to the higher speeds on landing, terrain following capabilities, and weapon delivery (as well as other factors), the pilot faced an increasingly higher load. Part of his/her time was taken up by switching monitoring between the extra-cockpit visual environment and the cockpit flight instruments. The HUD projects various flight parameters on a transparent glass so that the external environment can also be seen, eliminating the

pilot's need to look down into the cockpit to cross check the external environment with his/her flight instruments. Theoretically, the efficiency gained is a reduction in head movements, eye movements, and reaccommodations (Egan and Goodson, 1978). Intuitively, saving these wasted movements should enable the pilot to gather the information more quickly and thus more quickly process the information and respond accordingly. Although the HUD has been operational for many years, concerns still surround its use. These concerns include what information to include, how to display it, and how much information is needed.

Although general guidelines for visual displays exist (Lees, 1977; Rolfe and Chappelow, 1971) and more specific guidelines can be applied (e.g., Van Cott and Kinkade, 1972; Ketchel and Jenney, 1968), it seems that some of the HUD designs are not properly human-factored engineered (Egan and Goodson, 1978). Such general guidelines as: "Is the display in any way ambiguous? Does it take undue time to interpret? Is the indicated accuracy of the display adequate for or greater than is necessary for the achievement of the task objective? Do faults in the display become immediately apparent to the user without any possibility of misinterpretation?" which were offered by Rolfe and Chappelow (1971, p. 77) over ten years ago are violated in current HUD design. Part of the problem stems from the lack of empirical studies and the abundance of ambiguous terminology. An example of this is a report by Sperry Gyroscopes Co. (1963) which was reviewed in Egan and Goodson (1978). The report indicated the criteria "were that the information (i) enhanced instrument head-up

flight, (ii) enhanced visual head-up flight, (iii) improved the ability to assess partial information from the external world, (iv) was sampled frequently, or (v) improved IFR-VFR (instrument-visual flight rules) transition" (Egan and Goodson, 1978, p. 7). Although these guidelines are certainly worthwhile considering, a lack of precise language obscures the exact meaning for these criteria. In fact, the information required for a HUD during various mission profiles seems relatively unresearched (Egan and Goodson, 1978). These results obviously point out the need for additional research in the design of HUDs.

Not only is what information to display on a HUD an area of concern, but so is how to display the information. The symbology used to convey the information has become a source of confusion. This is attested to by the fact that different HUDs in different aircraft use different symbols. Although specifications exist as to the proper line width, brightness, etc., many of these specifications are extrapolated from other sources. Egan and Goodson (1978) point out that these values are "based on an educated guess of what the optimal values might be" (p. 13). A lack of empirical evidence exists as to the optimal values, and, in fact, feedback from some pilots found that seventy percent of them felt that the symbols interfered with night vision of the real world (Sheehan, 1972 in Egan and Goodson, 1978). The most alarming fact about this is that if pilots feel that it is more difficult to distinguish the real world with the HUD on, they will probably just turn it off and obtain their flight information

from the cockpit instrument displays. Thus the apparent advantages offered by the HUD will be wasted.

A third area that has been identified is the amount of information that should be displayed. Egan and Goodson (1978) report that display clutter is a complaint in every survey of pilots using HUDs. Although it is pointed out that display clutter is a poorly defined concept, it would be worthwhile to obtain the proper amount of information needed for the pilot to accomplish a task. Too much information tends to confuse the pilot and interfere with his/her ability to process the proper information. This again can be a major problem as is seen by Opittek's finding (1973; in Egan and Goodson, 1978) that 11 of 17 pilots turned off the HUD "at critical phases of a mission because it interfered with their performance" (p. 26). This can be seen as a major flaw in HUD design, because the HUD was designed to aid the pilot during critical phases of flight and, instead, it appears that the HUD interferes with performance just when it is needed most.

As can be seen, problems exist concerning what the optimal design is for the human-HUD relationship. Egan and Goodson (1978, p. 33) conclude that "there is very little hard evidence documenting the overall advantages of HUDs, and there is even less evidence concerning specific issues in the design of virtual-image displays." One would think that with the increase in technology, and the need to aid the already highly loaded pilot a great deal of research in this area would be conducted. Although Newman (1980) addresses some of the

operational problems associated with the use of HUDs, hard evidence is still lacking in this area.

The area of design and implementation of the HUD is an important one in need of more research. However, it is not the only aspect of the HUD that has been of concern. Although the physical attributes and structure of HUDs remain undefined (to some extent), a more general question with greater impact needs to be answered. The HUD was developed to aid the pilot under high loads. The simple question of whether the HUD accomplishes this is not easily answered. The previous sections on stress, workload, and information processing form a basis for analyzing the ability of the pilot to perform satisfactorily using the HUD. How these areas interact and are related to the symbology used will be discussed in the next section.

Possible Problems of HUD Use

The previous sections give us a framework from which we can begin to analyze how such variables as stress, information processing, workload, and symbology interact and affect human performance when using the HUD. Due to the fact that stress, information processing, and mental workload are so very interrelated any breakdown of these areas should be considered arbitrary. One should keep a system orientation when discussing these areas: a change in any one will affect the other two.

Stress

First, it must be remembered why the HUD was developed: to aid the pilot, especially under increased load conditions such as landings, weapon delivery, and terrain following. A report by Butterbaugh, Warner, Lovering, and Herron (1981) was designed to assess pilot workload. Although it was designed to be aircraft specific, several general conditions of high workload were identified. High workload was a result of crew station design (i.e., location of controls), in-flight procedures (checklists, communications, navigation), training, preparedness, and equipment malfunction. Another area that was associated with high workload was low-level flight profiles. Obviously there is a real need to decrease or at least aid the pilot during these high workload phases of flight, and thus the HUD seems to be an important factor. Secondly, and more importantly, during the high workload due to low-level flight HUDs do not seem to be aiding the pilot (as most aircraft capable of low-level flight profiles are equipped with HUDs), or if they are aiding the pilot they can at least use some improvement. Part of this apparent ineffectiveness may be due to the reason discussed earlier -- that the HUDs are frequently turned off because of their interference and this in itself is a major problem. If this is true, Goldstein and Dorfman (1975) have shown that under low load (one display) speed had little effect on performance, but under high load (three displays) increases in speed demands severely decreased performance. Thus if the pilot is within a low-level flight profile and must monitor several displays as well as the external environment, his/her poor performance may be due to this interaction of speed and load stress.

It has been shown that stressors cause a wide variety of physiological effects on the body. Although many biochemical substances are known to be produced or increased, little is known on the specific effects of them on the operator. In terms of chronic, long-term stress, these chemicals have been directly or indirectly linked to various psychosomatic illnesses. In the short term, these stressors affect performance in another way. Both Welford (1978) and Broadbent (1971) found that stress causes a reduction in the information perceived. When given a display it seems that some information is subjectively valued as less important and thus it is not given the amount of attention or effort given to the other information in the display. These results indicate that the type of information is an important consideration. In times of high load, such as terrain following, this load shedding or filtering phenomenon could play a critical role in the performance of the mission.

The theory proposed by Hamilton (1975) could, also, affect performance using the HUD. Given that the anxiety related cognitive data is to be avoided, then when similar stimuli are presented, the associated cognitive data may indeed take up part of the processing capability of the individual. Even if it does not take up some processing capability, it may very well interfere in other ways such as through distraction or automatic enabling of negative feelings.

Information processing

Both the single-channel, limited-capacity information processing model (which is similar to controlled processing) and the automatic processing model can play a large role in performance using the HUD. Given the dynamic nature of aircraft flight, the limited-capacity, single-channel model may well be appropriate when given a fairly inexperienced operator. Automatic processing is probably used when controlling the highly repetitious, redundant and often practiced elements of flight. These two modes, controlled processing and automatic processing, may have an effect which can be seen with different levels of operator experience. Crosby and Parkinson (1979) and Brainbridge (1978) found a difference between the performance of experienced and inexperienced operators due to what seemed to be an automatized type of behavior. As mentioned previously, the automatic processing ability takes a great deal of training to develop and is less flexible than controlled processing which takes more of the operator's time and attention. Thus, the development and use of automatic processing may be both a blessing and a curse. It is a blessing in the sense that it would help alleviate the load of continually interacting with tasks that are highly redundant and practiced and therefore not needing "conscious" control. However, in terms of cross-training or when the same stimulus demands a different response, automatic processing can greatly interfere with acquiring the new skills needed. Since many of the HUDs used in different aircraft are not standardized, this may indeed be a realistic problem.

Also, the amount of stress experienced may have a greatly varied effect on the experienced and inexperienced operator. The inexperienced operator has a much greater amount of mental work (or at least a feeling as though s/he does) than the experienced one, given the same task and situation (Bainbridge, 1978). Because of this, what seems to be low stress to the already experienced operator may well be medium or high stress for the inexperienced operator. This situation is compounded by the fact that stress tolerances and abilities are highly individualistic. Thus for an experienced operator a HUD may not be as stressful or demanding as for an inexperienced one. If processing of HUD-presented information were automatic as opposed to controlled, it would be interesting to know at what point does automatic processing occur? Much remains to be studied concerning the impact of Schneider and Shiffrin's (1977) model on the pilot's environment.

Mental Workload

Although mental workload can be discussed as a separate entity, it is easily seen how intricately related it is to these other areas. An integral component of mental workload is the feeling of how much work is being done. This area can be tapped through subjective measures such as rating scales and interviews. Its importance lies in the fact that not only would we be able to discern whether the HUD allows for better performance, but also the amount of information the operators feel is necessary to accomplish the task. If feedback from the pilots determines that the HUD interferes with, rather than

enhances, flight performance then some consideration should be given to the need for continuing its use or improving its design. If the HUD is found to enhance flying performance, then the amount of information needed to be displayed for various phases of flight should be determined to help eliminate unnecessary display clutter. A point needed to be added here is that the amount of workload experienced at any given point in time is a function of not only the present task but also of the past and future tasks as well. The pilot is monitoring aircraft systems such as fuel, planning future actions, and may be critiquing or processing past actions at any given point. Therefore the amount of workload may be more dynamic, based on the total experience of the operator, than previously alluded to.

Symbology

Another area of difficulty concerns the use of abstract symbology itself. It has been shown that the presentation of information must be similar in structure to the mental image the operator has and take into account the different operations affected. Without this similarity, the operator resorts to a supplementary coding mechanism which is generally a source of error (Ochanine, 1966 in Leplat and Pailhous, 1971). Thus if the symbology used in the HUD does not adequately reflect the proper relationship, then an additional processing level may be necessary and be the cause for an increased number of errors. Another area that is of concern has been studied by Bertera (1982). Bertera found that under stress subjects tend to process information at a more concrete level, suggesting that their

"abstraction ability" is reduced. An earlier study by Beier (in Cowen, 1952) also supports the finding that abstraction ability is reduced under stress. The use of abstract symbology in HUDs, then, may well be counterindicated during high loads due to the increase in supplementary coding and the tendency of the operator to process more concretely under stress.

Therefore a great deal of research (or rather lack of it) has left many questions regarding the use of HUDs unanswered. At the Ergonomics Society Annual Conference (August 1981) the implication of advanced systems as they apply to aviation was discussed. It was pointed out during the conference that the ability of head-up displays (as well as other advanced displays) to "really decrease the pilot's workload and enhance capability was questionable" (Adrian Harding in Taylor, 1981, p. 5). Other authors concur with this viewpoint (Egan and Goodson, 1978) and point out the need for further research in this area.

CHAPTER II

RESEARCH DESIGN AND METHOD

Research Design

Purpose

This research was designed to obtain data on the ability of the subject to process information displayed in different formats under a loading condition. The design was modelled after the present-day use of HUDs. Since different HUDs use different methods to display the same flight parameters, the ability of the human operator to process the information using different display methods was examined. The problem studied, then, was: Do the methods used to present various flight parameters in present-day HUDs result in differential human performance? To this end, two different presentation methods were selected and have been labelled as either concrete or abstract. Although the terms, concrete and abstract, carry with them an intuitive feeling for each presentation method, they require further definition to understand their scope and limitations as used in this study.

Definition

To compare the terms, concrete and abstract, imagine a forced choice experiment. We are asked to label two paintings as either concrete or abstract. If the two paintings were a Renoir and a later Picasso, which painting would be labelled as concrete? Intuitively,

we would probably label the Renoir concrete while the Picasso would probably be labelled abstract. Why does this seem likely? They both are physical entities and painted in the same medium; they both may contain the same number bits of information; they both are visual stimuli; they both are representations of the same subject; etc. Wherein lies the difference? The difference may be in our ability to form a one-to-one relationship between what is represented on the canvas and what is found in the real world. Not only would the sum of the parts be a necessary component of the overall concreteness or abstractness of the painting, but so would the relative relationships between the parts.

This representational paradigm has been used to study the difference between concrete representations in memory and abstract representations in memory. Spoehr and Lehmkukle (1982) discuss several experiments that deal with the processing of abstract and concrete words. The paradigm used in this research was a paired-associate learning task that Paivio has studied extensively. According to Paivio (in Spoehr and Lehmkukle, 1982), the learning of a paired association is enhanced by the presence of a "conceptual peg" on which we can hang associative links of a paired word. Imagery seems to be a good conceptual peg and the difference between concrete and abstract nouns may be the difference in the ability to form a mental, representational image. Thus a concrete word such as "apple" is more easily represented mentally than an abstract word such as "truth". However, confounding factors such as meaningfulness and familiarity have also been identified as influencing the ability to

label one word as concrete and another as abstract, with the more concrete word being more meaningful and familiar. When controlling for these two factors, it seems that concrete and abstract nouns differ in the type of processing available to determine their meanings. Paivio has suggested that there are two modes available to us to process the nouns; we can process the nouns through a verbal (linguistic) mode or through a nonverbal (imaginal) mode. Due to the fact that our experiences with abstract nouns are based solely on linguistic experiences, abstract nouns are processed verbally. We are unable to physically encounter such abstract concepts as truth, honor, soul, etc. and therefore must process them through their verbal meaning only. Concrete words, however, can be processed verbally or nonverbally because we have experienced them both verbally or semantically and physically or sensorially. This processing may also be an integrated one (Marshark and Paivio, 1977). We may process abstract and concrete concepts based not only on their semantic meanings but on the relationships derived from them based on the context and our knowledge of language and the world. From this we may form an integrated mental representation. Perhaps the reason that a Renoir may be labelled concrete is that it more accurately reflects our mental representation of the object portrayed.

For this research, we will use a similar construct to define our abstract and concrete displays. However, since each display will be providing information on the same flight parameters, it is their relative abstractness that concerns us. Each HUD will display six parameters: airspeed, altitude, vertical velocity, heading, bank

angle, and pitch. On one display the airspeed, altitude, and vertical velocity will be displayed as tapes with moving pointers indicating the various values. This display is the relatively concrete HUD. For the relatively abstract HUD, all parameters except bank angle and pitch will be presented in digital form.

Although the same concepts are being displayed, the relatively concrete and abstract HUDs can be so labelled based on the method of presenting the parameters. The flight parameters themselves can be considered as abstract concepts. Airspeed, for example, is not a physical, tangible object: we would be hard-pressed to mentally imagine an "airspeed". What is available to us, though, is the measured value of an airspeed. The concept of airspeed includes a specified quantity or amount: we can have more or less airspeed or we can increase or decrease airspeed. In this way the measured value of airspeed becomes the concrete representation of the abstract concept. An airspeed of 125 mph is faster than an airspeed of 120 mph; it is not only faster but we know by how much -- 5 mph; yet, we cannot picture an "airspeed" -- only its measurement. Altitude and vertical velocity can be thought of in the same way. We can go up or down, increase or decrease, etc.; yet, we can hardly picture an altitude or vertical velocity. These familiar abstract concepts are easily understood because of the contact we have had with their measurements. We have encountered such displays in many aspects of daily living: rulers, speedometers, thermometers, clocks, etc. All these displays are representations of familiar concepts that can be thought of as quite abstract. However we may use the term concrete for these

concepts because of the representational nature of the displays (up or down, increase or decrease, more or less, etc.) and the great amount of familiarity we have with them. The method of measurement has "concretized" the concept and allowed us to form a mental representation of what are in reality abstract concepts. In this way, airspeed, altitude, and vertical velocity will be presented on the relatively concrete display as tapes with moving pointers that have the representation of increasing or decreasing, up or down, more or less, etc. This presentation style enhances the familiar mental representations of these concepts that we have encountered previously (i.e. as speed in automobile speedometers). Then for the more abstract display we use digits to present the flight parameters. This display can be labelled as relatively abstract because the method of presentation does not enhance a mental representation of these variables. The subject will have to supply the mental representations from the various values given. Just as in concrete/abstract words where the analog or mental representation of the concrete word is more apparent than the abstract word, in the tape (relatively concrete)/digit (relatively abstract) displays the analog or mental representation of the tape-displayed parameters is more apparent than the digit-displayed parameters.

Method

Subjects

A total of 36 flight-naive subjects (12 female) began the study. All subjects were volunteers who responded to a flyer placed in the

Industrial Engineering/Operations Research Department or Department of Aerospace Studies at the University of Massachusetts. Although several subjects did not come from one of these areas, the majority of subjects did. Of the original 36 subjects, 19 subjects (5 female) completed all aspects of the experiment. The mean age of these subjects was 21.47 years with a range from 18 to 25 years. Of the remaining 17 subjects, 12 completed some aspects of the experiment enabling data from them to be incorporated in the analysis. Data from the remaining five subjects could not be included in the data analysis due to their incompleteness. There are two primary reasons for this attrition rate. Time was the primary reason for the lack of complete data on the 12 subjects who partially completed the experiment. The experiment was conducted during the latter half of the school's second semester. Due to the unanticipated length of training encountered, these subjects were unable to complete the study prior to the end of the semester and their departure from the campus. The remaining five subjects withdrew from the experiment without informing the experimenter as to their reasons. The implications of this attrition rate will be addressed in the discussion section.

Apparatus

Three different methods of displaying the flight parameters were used. In addition to the concrete and abstract HUDs described previously, the flight parameters were displayed on a PACER MKII desktop flight simulator. The Pacer represents the general class of single-engine, light aircraft and includes a full instrument panel

(see Figure 4). The two HUDs were displayed on an AMDEX 13-inch diagonal color I monitor placed level with the top of the Pacer simulator, a position similar to the location of a HUD in an aircraft. The specific HUD displays are shown in Figures 5A and B. The Pacer simulator was connected to an Apple II plus computer which digitized the signals and fed them to the AMDEX monitor to drive the simulated HUDs (Figure 6).

A secondary task was developed following Brown (1962, 1965). This task was an auditory task which presented a series of random digits (1 through 9) at a rate of one every 1.25 seconds. The series of digits were taken from a table of random numbers with the following constraints: (1) that no digit occurs twice in succession; (2) that an odd-even-odd sequence occurs at least once every 30 seconds; and (3) that no embedded sequences occur (i.e. odd-even-odd-even-odd). The subjects' task was to verbally respond "NOW" immediately after every odd-even-odd sequence.

Procedure

The original 36 subjects were randomly assigned to one of six groups (with the exception that at least two female subjects appear in each group). The groups were counterbalanced across the order of display presentation: Pacer, concrete HUD, and abstract HUD. The first session was an introductory session. A 30 minute (approximately) cassette tape was used to introduce the subject to his/her first display. The tape explained each of the instruments and their relationship to each other, and also included a description of

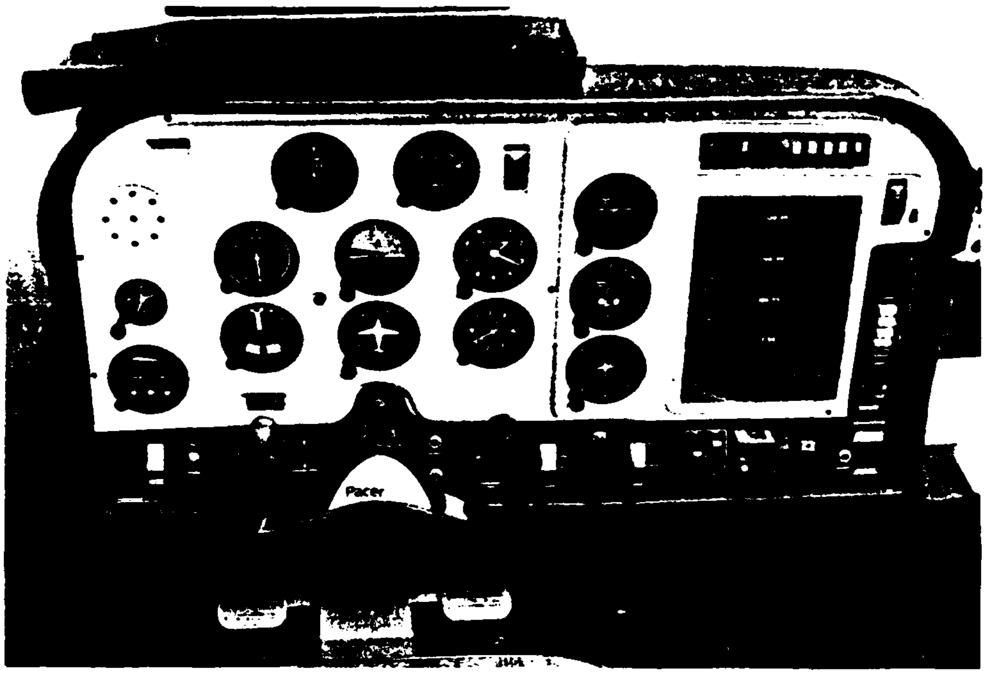
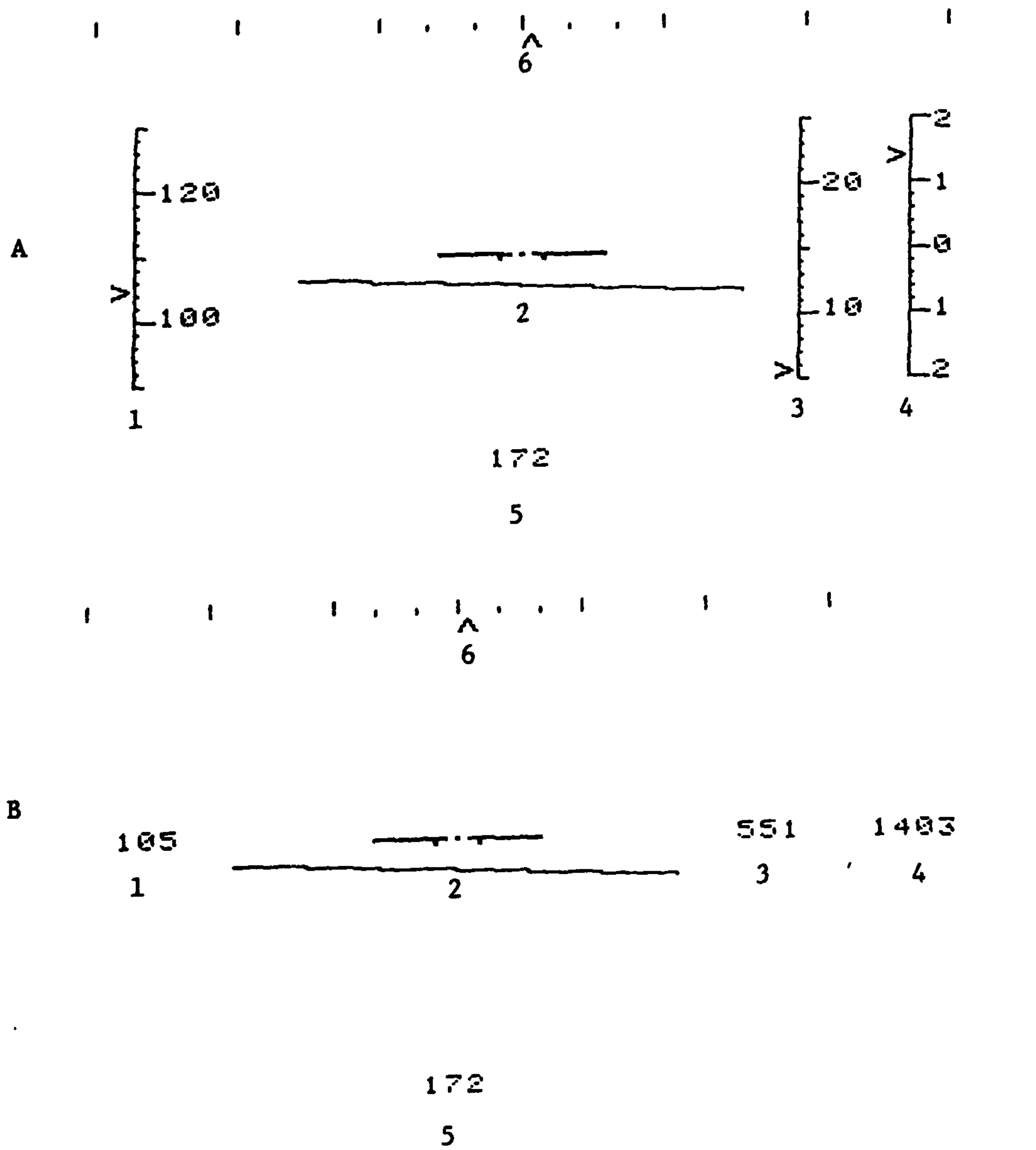


Figure 4. Pacer MKII Desktop Simulator



1 - Airspeed (mph)

2 - Artificial horizon (pitch and roll)

3 - Altitude (feet)

4 - Vertical velocity (feet/minute)

5 - Heading

6 - Bank angle

Figure 5. Simulated HUDs:
A. Relatively concrete HUD;
B. Relatively abstract HUD

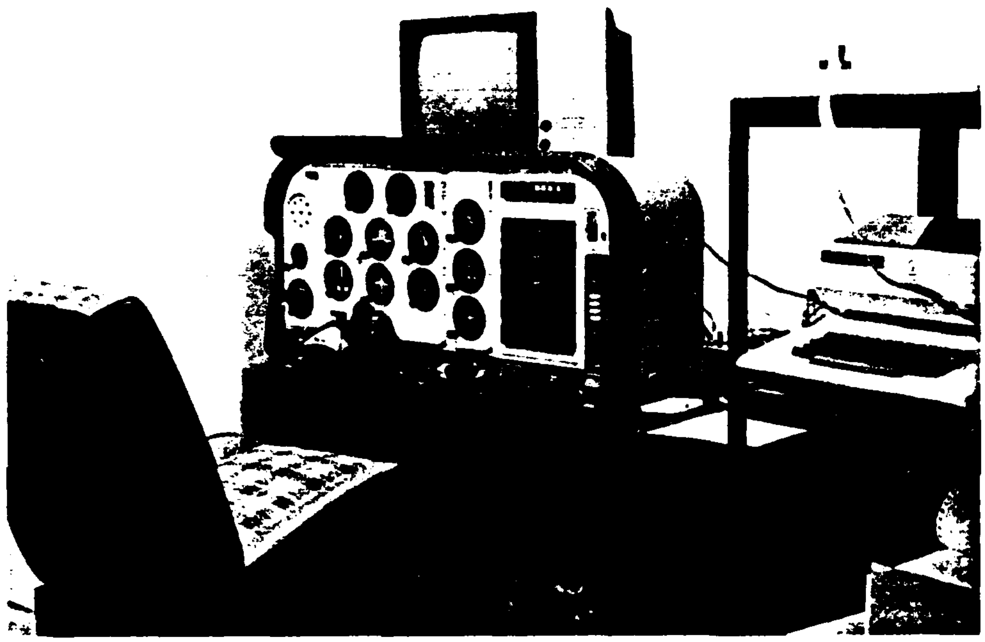


Figure 6. Experimental apparatus showing the Pacer MKII desktop simulator and the AMDEX color I monitor

the five flight maneuvers that the subject would learn. This introductory session also provided the opportunity for the subject to practice each of the maneuvers. The five flight maneuvers and their target values were: (1) to climb 1000 feet at 500 feet per minute at 115 mph; (2) to fly straight and level for two minutes at 115 mph; (3) to descend 1000 feet at 500 feet per minute at 115 mph; (4) and (5) to turn right and left, respectively, at 30 degrees of bank for 180 degrees at 115 mph and zero vertical velocity.

Training

The maneuvers were to be performed in the order above while on turbulence level one until the successful completion of each of the maneuvers. The slight turbulence was used in order to keep the subject an active controller during the flight maneuvers. The successful completion of each maneuver required the subject to keep the flight parameters within the acceptable limits shown in Table 1 (derived from Koonce and McCloy, 1981; Koonce and Berry, 1980).

TABLE 1

	Maneuver				
Parameters	Climb	Cruise	Descent	Rt. Turn	Lt. Turn
Vert. vel. (ft/minute)	500'±175'	0±75'	500'±175	0±150'	0±150'
Bank angle (degrees)	--	--	--	30±8	30±8
Airspeed (mph)	115±4	115±2	115±4	115±4	115±4
Altitude	2000'±60' at level off	2000'±40'	1000'±60 at level off	1000±75'	1000±75'
Heading (degrees)	±8	±3	±8	±7 on rollout	±7 on rollout

If a subject failed to perform a maneuver within the acceptable limits then that maneuver was not scored, and the subject continued through the prescribed order for the maneuvers. The subject then completed another set of maneuvers deleting any maneuver that s/he successfully completed previously. In this way the order of maneuvers was preserved. The flying task is modelled after Koonce and his colleagues (Koonce and McCloy, 1981; Koonce and Berry, 1980). The selection of these maneuvers was based on two reasons. These flight maneuvers form the fundamental skills necessary for any flying task, and the proper accomplishment of each maneuver (except for the straight and level cruise) requires the change in only one flight parameter. Climb and descent (once established) require a change in

altitude, the turns require a change in heading. After the subjects completed their first display, they were introduced to their next display and the instruments were explained. The subjects then completed the same process for all maneuvers on the second and third displays. The acceptable limits on each of the maneuvers on each display remained the same. The training phase ended with the successful completion of all maneuvers on each of the displays.

Testing

After the subjects could perform all five maneuvers within the preset criteria, they were introduced to the auditory task. The subjects read a brief, written instruction on the auditory task and their questions were answered. They then had a five minute practice trial on the task. Each subject was allowed six mistakes out of a possible 30 (approximately) correct sequences before s/he was considered to have failed the auditory task introduction. After a failure on the auditory task introduction, a subject was permitted two more tries during the same session, otherwise the subject completed the secondary task introduction at the following session. The number of trials on the secondary task introductions ranged from one trial to four trials with a mean of 1.474 trials. Only three of the 22 subjects that attempted the secondary task took more than two trials to successfully complete the secondary task introduction.

Following the introduction to the secondary task, the subject read the instructions for the testing session which read in part: "Your first priority during this task is to fly the maneuvers as best

you can, try not to let the auditory task interfere with your flying performance." This established the priorities that the experimenter wished the subject to form when accomplishing both the primary, flying task and secondary, auditory task simultaneously. The subjects were then asked to fly the five maneuvers with each of the displays in the order that the displays were learned while listening to the auditory task. The same flight criteria as in the training sessions were required for a successful completion of a testing trial. There were no criteria for successful auditory task completion.

Measures

In addition to the trials to criterion on the training phase (TRAIN) (minimum of five, one for each maneuver), the trials to criterion on the testing phase (TEST) was also obtained. The percent correct digit sequences deleted on both the first trial during the testing phase (FIRST) and the last trial during the testing phase (LAST) were recorded. The subject need not have passed the maneuver in order to obtain the FIRST score, however, the LAST score was only based on passed maneuvers. Therefore, the FIRST and LAST scores would be the same when the subject passed that particular maneuver on his/her first try.

After the subjects completed the testing phase, they were administered a questionnaire (see Appendix C). The questionnaire was designed to obtain the subjects' feelings toward the various aspects of the experiment.

CHAPTER III

RESULTS AND DISCUSSION

Results

The following results are presented in shortened form; the full ANOVA tables are presented in Appendix A. In general the model used was the repeated measures model. The use of this model grants us more efficiency and power for the given number of subjects. Separate analyses were carried out and are presented below.

Training and testing analysis

Display by phase by subjects. This two-factored repeated measures design allowed us to look at the main effect of display (Pacer, Concrete HUD, and Abstract HUD), Phase (trials to criterion of the training phase, TRAIN, and on the testing phase, TEST), and the interaction between the two (Display x Phase). There was a very significant effect due to the interaction term, $F_{2,36} = 147.57$ ($p = .0000079$; see Table 2). This makes interpretation of the main effects somewhat difficult, since the performance on a particular display changes as a factor of time. However I will present their F values for the sake of completeness: The display main effect had an F value of $F_{2,36} = 2.76$($p = .0764$) and the F value for the phase main effect was $F_{1,18} = 157.464$ ($p = .00003$). I decided to follow this analysis

with separate analyses for displays on each phase (training and testing).

Order effect in learning the displays. First, it seemed likely that there would be an effect due to which display the subject learned first, second or third. To analyze this, One-way ANOVAs were performed on each order of presentation over the three displays. For the first display that the subjects learned there were no significant differences between the mean trials to criterion across the three displays: Pacer, concrete HUD, or abstract HUD ($F_{2,28} = 2.87$, $p = .0735$). This result was also found for the subjects' mean trials to criterion on the second and third display ($F_{2,25} = 3.188$, $p = .0584$; $F_{2,23} = 1.611$, $p = .2214$ respectively). Therefore, it seems that within a particular order of presentation there is no significant difference on the trials to criterion between the Pacer, concrete HUD, and abstract HUD (see Tables 3 through 5).

Training phase (TRAIN) by subjects. Two choices were available with this analysis. The first was to include only the subjects that completed the entire study in the analysis (n = 19) and the second was to include all subjects that completed at least the training phase (n = 26). I did both. When looking at only the subjects that completed the entire experimen., there was no significant difference between the mean trials to criterion across the three displays ($F_{2,36} = 1.669$, $p = .2027$). However when all the subjects that completed the training phase were included in the analysis, the results were very close to

significance ($F_{2,50} = 3.125$, $p = .0526$). A Scheffe test was performed and it was the difference between the Pacer and Abstract HUD which seemed to be responsible (critical difference$_{.0526}$ = 6.786, obtained difference = 6.7693; see Table 6). The implications for these results will be presented in the discussion section.

Testing by subjects. The trials to criterion on the test phase were then analyzed. An $F_{2,36} = 3.151$ was obtained and the probability was .0548 (Table 7) which is again very close to significance. Again a Scheffe test was performed. Although there were no significance differences between the means at the .05 level there was a significant difference between the Pacer and both the Concrete HUD and Abstract HUD at the .10 level. These results are illustrated in Figure 7 for the 19 subject case.

Percent correct digit sequences

Percent correct sequences on first trial (FIRST by subjects). The mean number of digit sequences correctly identified by the subjects on their first trial did not significantly differ across displays ($F_{2,36} = 1.5754$; $p = .2209$, Table 8).

Percent correct sequences on last trial (LAST by subjects). Where the FIRST by subjects did not show any significant difference, the LAST by subjects certainly did ($F_{2,36} = 3.8361$; $p = .0309$, Table

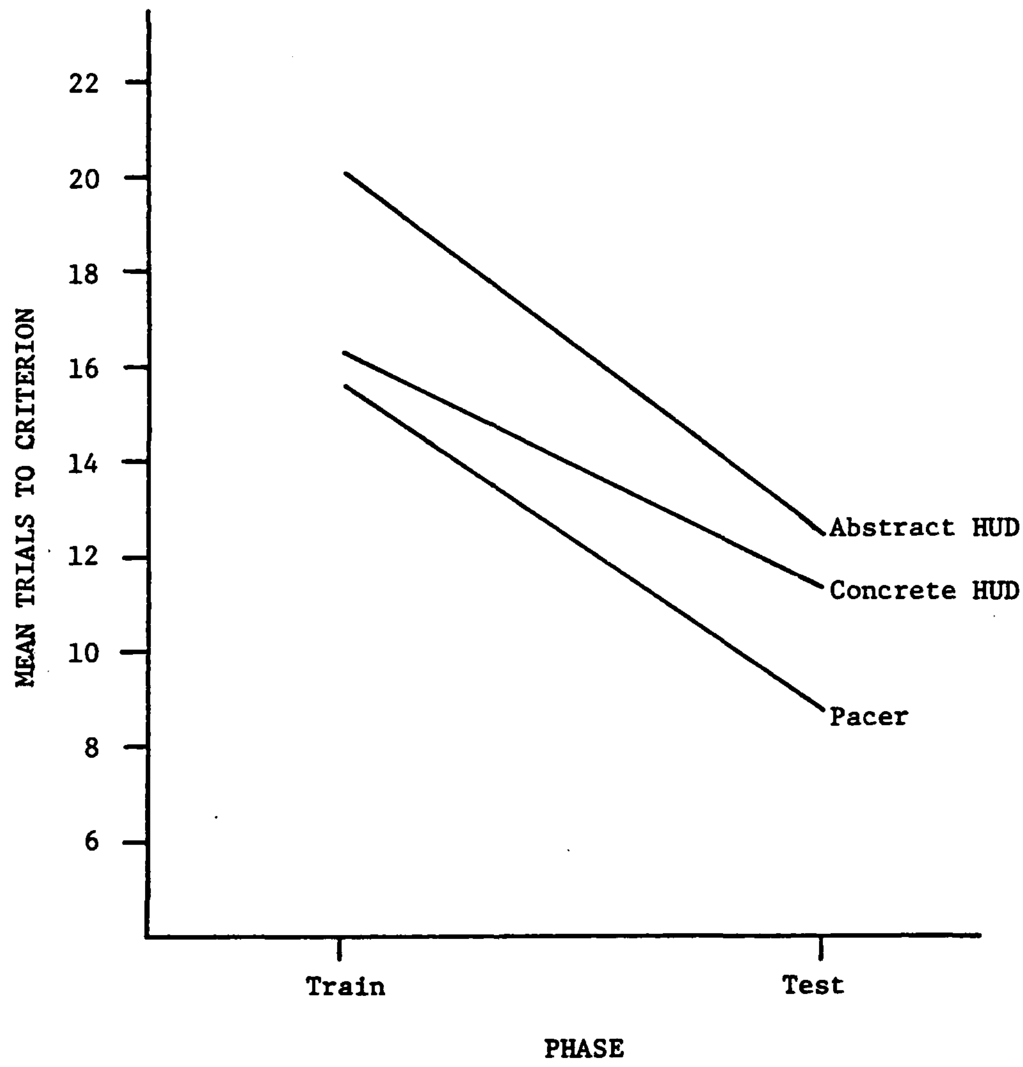


Figure 7. Mean trials to criterion on training and testing phase.

9). Figure 8 illustrates these results. Although there was no significant difference between the Concrete and Abstract HUDs, the mean percent correct digit sequences was significantly higher using the Pacer than all combinations of the HUDs by a post hoc analysis.

Questionnaire analyses

An end of experiment questionnaire was administered to the subjects (see Appendix C). This questionnaire was designed to tap the subjects' subjective ratings of various aspects of the experiment. The first question was to determine how difficult the subject felt the displays were to learn, disregarding the order that they were presented. Figure 9 shows a general tendency for the subjects to feel that the Pacer was the easiest display followed by the Concrete HUD. The Abstract HUD was rated as the most difficult. These results paralleled what was found in the analysis of the training phase.

The subjects' ratings of workload under the displays also followed this general pattern as can be seen in Figure 10. This data was taken from the subjects' responses on question 3 of the questionnaire. The wording was left somewhat ambiguous by intention to allow the subject to rate his/her workload by any measure he/she desired. However, the general tendency of increasing difficulty from Pacer to Abstract HUD still seems to hold.

Question 4 was included to determine how the addition of the secondary task changed the subjects' perception of workload. A chi-square test was performed on Questions 3 and 4, and the results suggest that the addition of the secondary task resulted in an

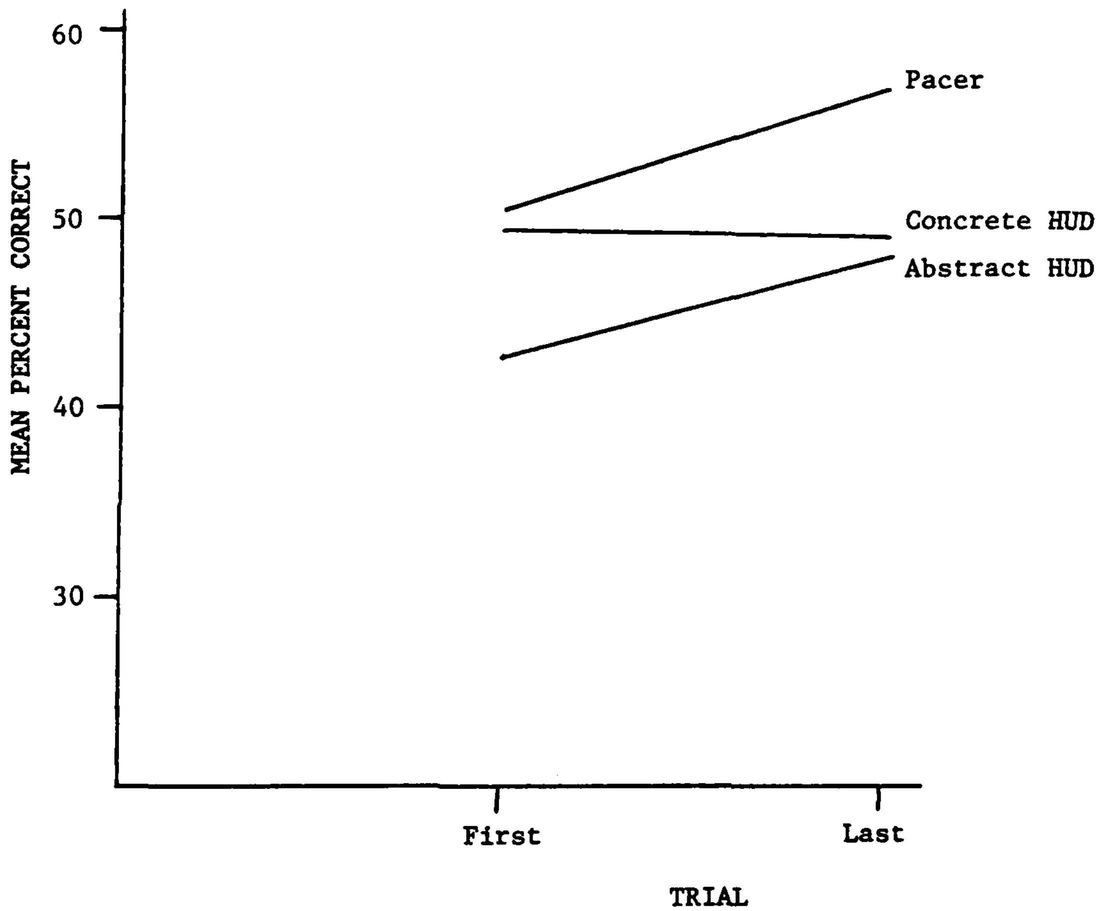


Figure 8. Mean percent correct digit sequences on first and last trial.

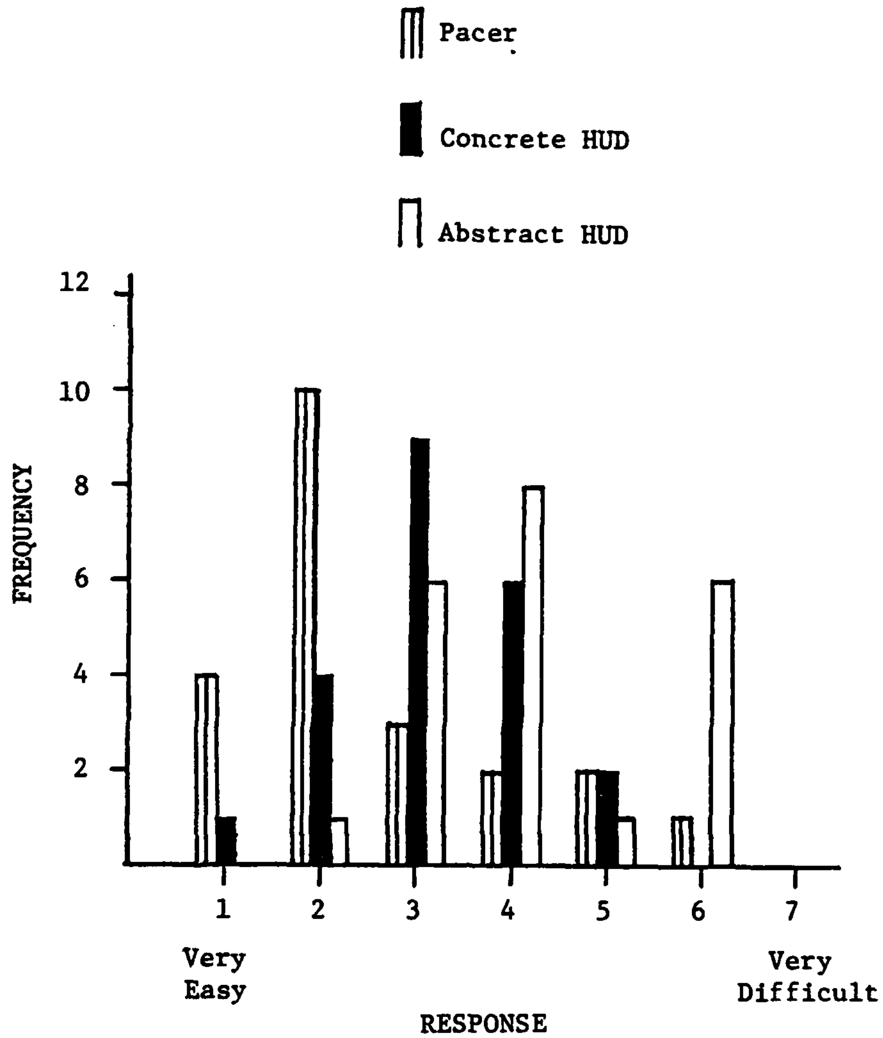


Figure 9. Frequency of responses to question #1: "How difficult was each of the displays to learn?"

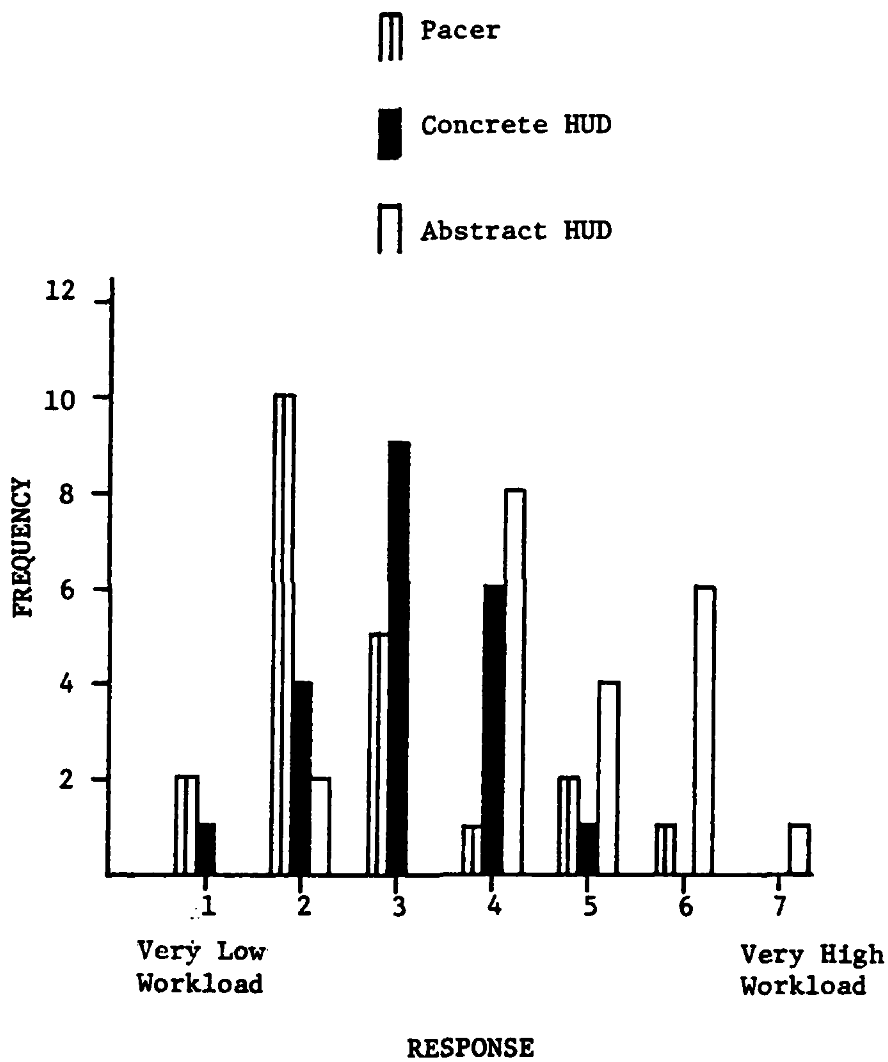


Figure 10. Frequency of responses to question #3: "Without the auditory task and averaged over all maneuvers, how would you rate the three displays based on the amount of workload you felt in flying to the preset criteria?"

increased feeling of workload, with the tendencies addressed above still holding. This is reassuring since we would have liked to have determined if the secondary task was in fact felt to be a loading condition (Figure 11).

Discussion

The high attrition rate of the subjects had a significant impact on this study. It seems that the poorer subjects were "selected-out" leaving only the better performers. This is seen by the TRAIN by Subjects analyses. When all the subjects that completed the training phase were included, a significant difference between learning the displays was apparent, however, when only the subjects that completed the study were selected, no difference was found. If given more time, it is still unclear whether all subjects that completed part of the study would have dropped out due to the increasing in frustration, weakening in motivation, and decreasing in novelty of the situation. These factors could also have "selected-out" the poorer performers even given unlimited time.

Although each of the specific displays was unfamiliar to the flight-naive subjects, it seems that the displays themselves were inherently difficult. Although the Pacer seemed to be easier than the HUDs to learn, no difference was found between the two HUDs. The subjects may have been more familiar with the type of displays found on the Pacer which are similar to common gauges found in automobiles and other systems. The HUDs, however, presented information

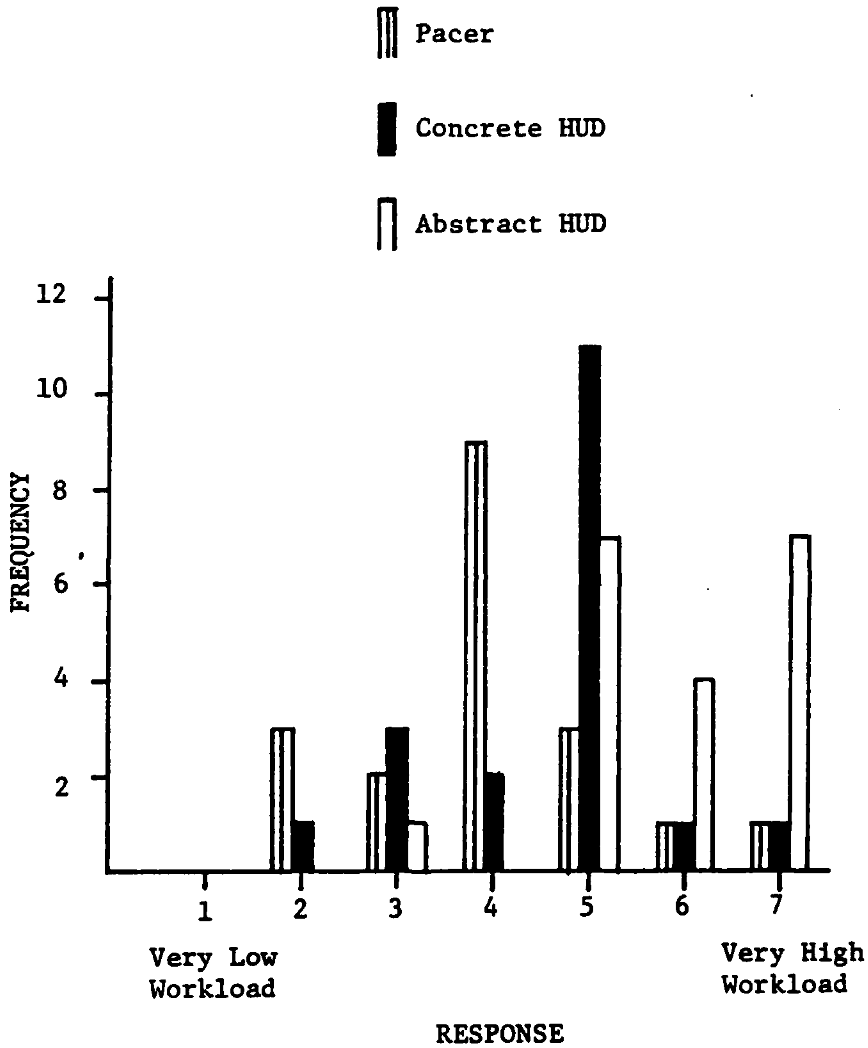


Figure 11. Frequencies of responses to question #4: "With respect to the auditory task, how would you rate the following in terms of the amount of workload you experienced?
Auditory task and: (display)"

differently than most of us are accustomed to using. Therefore, the displays may not have been equally unfamiliar.

Another factor may have been the sensitivity of the simulated displays. Due to the fact that the two HUDs were being driven by the Pacer through a computer, there was a longer response time between control input and displayed response for the two HUDs than that found with the Pacer. This may have contributed greatly to the difficulty of the two HUDs. It is well established that an increase in feedback time decreases performance (see Rouse, 1980 for a discussion of this). It may be the sensitivity of the display rather than the style of information presented on the display that is responsible for the differences found in learning the flying task.

The trials to criterion for the testing phase was also marginally significant. It seemed that the ability of the subject to simultaneously perform both the flying and auditory task was dependent on the display. Although by requiring all the subjects to perform to the specified criterion levels, no difference on the trials to criterion on the testing phase was expected. There may have been some "lucky passes" where the subject really did not pass a particular maneuver but was passed nonetheless or "unlucky failures" where the subject just missed passing the maneuver. This introduces a possibility of scoring errors.

No significant findings of the percent correct digit sequences on the first trial was found. Since the subject may not have passed the maneuvers on these trials, s/he may not have established the proper priority for the primary and secondary task. Another reason for

looking at this analysis was to see if there was a change in strategy between the trials where the subjects did not pass and those where they did. The significant LAST by subjects effects suggest that one of these may be taking place. The subject may have given the secondary task too high a priority and therefore needed to change the strategy being used. This significant effect also seems to show that although no difference in learning the displays exists (comparing only subjects that completed the experiment), once learned and put under a loading task there was a difference in performing the secondary task. Since the primary task by definition had to have been passed, it seems that the amount of workload experienced with each of the displays as measured by the secondary task changed. Although we cannot, at this point, tell if it is due to the presentation style of the information on the displays or the differences in sensitivity, we can say that given no differences in learning the displays, a difference does surface when performing the task under a loading stress.

Although originally the study was designed to be a fully counterbalanced, repeated measure design, the high atrtrition rate encountered made this unobtainable. For a more detailed analysis of the number of subjects in each group and the order of display presentation please refer to Appendix B.

Observations

An observation that cannot be tested with the results of this study but should be noted for future studies is the strategy that seemed to be used while flying the primary and secondary task

concurrently. The flying task was assigned the primary priority and because of this a time-sharing strategy was developed. Whenever a maneuver required a change in the beginning or end of the maneuver (all except cruise) the subjects tended to ignore the secondary task until the maneuver was stabilized. This cannot be tested with the present data because of the random appearance of the odd-even-odd sequences. Therefore on some maneuvers no sequences appeared during this transition phase, whereas on others, several sequences may have occurred. If the number of sequences was constant for all phases of a maneuver the time-sharing strategy may be analyzed. This strategy was apparent to the experimenter, as well as to the subjects themselves for several subjects commented that this was the strategy employed.

With respect to the questionnaire data, it seems that the results found in the analysis of the data were in agreement with the subjective ratings of the subject. The subjects felt that the abstract HUD was the hardest and they tended to do the poorest on it whereas the Pacer was felt to be the easiest and they tended to do the best with it. As mentioned earlier, subjective measures of workload seem to be a promising technique in the assessment of workload, even this crude subjective assessment seemed to discriminate workload differences.

Recommendations

I recommend that an additional study similar in design to this study be conducted with highly flight-experienced subjects. Due to their familiarity with the flight displays, it is hypothesized that the abstract HUD will even have a greater effect on performance.

References

Appley, M.H. and Trumball, R. On the concept of psychological stress. In W. C. Howell and I. L. Goldstein Engineering Psychology: Current Perspectives in Research. New York: Appleton-Century-Crofts, 1971.

Bainbridge, L. Forgotten alternatives in skill and workload. Ergonomics, 1978, 21 (3), 169-185.

Bertera, J. Personal Communication. Department of Psychology, University of Massachusetts, Amherst, MA, May 1982.

Broadbent, D. E. Decision and Stress. New York: Academic Press, 1971.

Brown, I. D. Spare mental capacity of car drivers. Ergonomics, 1962, 5, 247-250.

Brown, I. D. A comparison of tasks used to measure fatigue in car drivers. Ergonomics, 1965, 8, 467-473.

Butterbaugh, L., Warner, D., Lovering, P. and Herron, S. Pilot workload: A survey of operational problems. Wright-Patterson AFB, Ohio: Flight Dynamics Laboratory, AFWAL-TR-81-3011, August 1981.

Cowen, E. L. The influence of varying degress of psychological stress on problem-solving rigidity. Journal of Abnormal and Social Psychology, 1952, 47, 512-519.

Crosby, J. V. and Parkinson, S. R. A dual task investigation of pilots' skill level. Ergonomics, 1979, 22 (12), 1301-1313.

Egan, D. E. and Goodson, J. E. Human factors engineering for head-up displays: A review of military specifications and recommendations for research. Pensacola, Florida: Naval Aerospace Medical Research Laboratory, Monograph 23, April 1978.

Fisher, D. L. Limited-channel models of automatic detection: Capacity and scanning in visual search. Psychological Review, 1982, 89(6), 662-692.

Goldstein, I. L. and Dorfman, P. W. Speed and load stress as determinants of performance in a time sharing task. Human Factors, 1978, 20 (3), 603-609.

Hochberg, J. E. Perception (2nd Ed.). Englewood Cliffs: Prentice-Hall, Inc., 1978.

Hogan, R. and Hogan, J. C. Subjective correlates of stress and human performance. In Earl A. Alluisi and Edwin A. Fleishman (Ed.) Human Performance and Productivity: Stress and Performance Effectiveness. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers, 1982.

Ketchel, J. M. and Jenney, L. L. Electronic and optical generated aircraft displays: A study of standardization requirements, Washington, D.C.: Office of Naval Research, JANAIR. TR-680505, May 1968.

Knowles, W. B. Operator loading tasks. Human Factors, 1963, 5, 155-161.

Koonce, J. M. and McCloy, T. M. Sex as a moderator variable in selection and training of persons for learning flight maneuvers. First Symposium on Aviation Psychology. Columbus, Ohio: Ohio

State University Aviation Psychology Laboratory, April 1981, 261-166.

Koonce, J. M. and Berry, G. A. Comparison of males and females in the prediction of basic flight performance. Seventh Symposium on Psychology in the Department of Defense, Colorado Springs, Colorado: U.S. Air Force Academy, April 1980, 612-615.

Lachman, R; Lachman, J. L. and Butterfield, E. C. Cognitive Psychology and Information Processing: An Introduction. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers, 1979.

Lazarus, R. S. A laboratory approach to the dynamics of psychological stress. American Psychologist, 1964, 19, 400-411.

Lazarus, R. S., Deese, J. and Osler, S. F. The effects of psychological stress upon performaance. Psychological Bulletin, 1952, 49, 293-317.

Lees, R. The man/machine relationship. Displays for Man Machine Systems, 4-7 April 1977.

Leplat, J. and Pailhous, J. The analysis and evaluation of mental work. In W. T. Singleton, J. G. Fox, and D. Whitfield (Eds.) Measurement of Man at Work. London: Taylor & Francis LTD., 1971.

Marshark, M. and Paivio, A. Integrative processing of concrete and abstract sentences. Journal of Verbal Learning and Verbal Behavior, 1977, 16, 217-231.

Moray, N., Johanssen, G., Pew, R., Rasmussen, J., Sanders, A. and Wickens, C. Final report of experimental psychology group. In

N. Moray (Ed.) Mental Workload: Its Theory and Measurements. New York: Plenum Press, 1979.

Mulder, G. Mental load, mental effort and attention. In N. Moray (Ed.) Mental Workload: Its Theory and Measurement. New York: Plenum Press, 1979.

Newman, R. L. Operational problems associated with Head-Up Displays during instrument flight. Wright-Patterson AFB, Ohio: Air Force Aerospace Medical Research Laboratory, Technical Report AFAMRL-TR-80-116, 1980.

Ogden, G. D., Levine, J. M., and Eisner, E. J. Measurement of workload by secondary tasks. Human Factors, 1979, 21 (5), 529-548.

Rolfe, J. M. and Chappelow, J. W. Display of information in the aircraft cockpit. Displays, 7-10 September 1971, 77.

Ryan, C. Reassessing the automaticity-control distinction: Item recognition as a paradigm case. Psychological Review, 1983, 90(2), 171-178.

Sanders, A. F. Some remarks on mental load. In N. Moray (Ed.) Mental Workload: Its Theory and Measurement. New York: Plenum Press, 1979.

Schaffer, H. R. Behavior under stress: A neurophysiological hypothesis. Psychological Review, 1954, 61, 323-333.

Schneider, W. and Shiffrin, R. M. Controlled and automatic human information processing: I Detection, search, and attention. Psychological Review, 1977, 84 (1), 1-66.

Schuler, R. S. Definition and conceptualization of stress in organizations. Organizational Behavior and Human Performance, 1980, 25, 184-215.

Selye, H. Stress Without Distress, New York: J. B. Lippincott Company, 1974.

Shiffrin, R. M. and Schneider, W. Controlled and automatic human information processing: II Perceptual learning, automatic attending and a general theory. Psychological Review, 1977, 84 (2), 127-190.

Spoehr, K. T. and Lehmkukle, S. W. Visual Information Processing. San Francisco: W. H. Freeman and Company, 1982.

Taylor, R. M. Aviation ergonomics -- The implications of advanced systems. The Ergonomist, August 1981, 4-5.

Thompson, R. F. Introduction to Physiological Psychology. New York: Harper & Row, 1975.

Ursin, H., Baade, E. and Levine, S. (Eds.) Psychobiology of Stress: A Study of Coping Man. New York: Academic Press, 1978.

Van Cott, H. P. and Kinkade, R. G. Human Engineering Guide to Equipment Design. Washington, D.C.: U. S. Government Printing Office, 1972.

Welford, A. T. Mental workload as a function of demand, capacity, strategy, and skill. Ergonomics, 1978, 21 (3), 151-167.

Welford, A. T. Stress and performance. In A. T. Welford (Ed.) Man Under Stress. London: Taylor & Francis Ltd., 1974.

Wierwille, W. W. Physiological measures of aircrew mental workload. Human Factors, 1979, 21 (5), 575-593.

Willeges, R. C. and Wierwille, W. W. Behavioral measures of aircrew mental workload. Human Factors, 1979, 21 (5), 549-574.

APPENDIX A

Table 2

Analysis of Variance of Trials to Criterion on Displays and Phase of Experiment (Training and Testing Phase)

Source	df	SS	MS	F	p
Total	113	6913.623	61.18		
Subjects	18	691.79	38.43		
Display	2	379.62	189.81	2.765	0.0764
Subj x Display	36	2471.21	68.64		
Phase	1	1286.74	1286.74	157.46	0.00003
Subj x Phase	18	147.09	8.17		
Display x Phase	2	1726.597	863.299	147.57	0.0000079
Subj x Disp x Phase	36	210.58	5.85		

Table 3

ANOVA of Trials to Criterion on Training of the First Display Only

Source	df	SS	MS	F	p
Total	30	4616.97			
Display	2	785.55	392.78	2.87	.05
Error	28	3831.42	136.84		

	Pacer	Concrete HUD	Abstract HUD
$\bar{X}$	21.62	27.29	33.09
n	13	7	11

Table 4

ANOVA of Trials to Criterion on Training of the Second Display Only

Source	df	SS	MS	F	p
Total	27	1600.7			
Display	2	325.28	162.64	3.18	.05
Error	25	1275.41	51.01		

	Pacer	Concrete HUD	Abstract HUD
$\bar{X}$	12.125	18.17	20.88
n	8	12	8

Table 5

ANOVA of Trials to Criterion on
Third Display Only

Source	df	SS	MS	F	p
Total	25	876.62	53.87	1.611	.05
Display	23	768.87	33.43		

	Pacer	Concrete HUD	Abstract HUD
$\bar{X}$	11.43	11.78	15.8
n	7	9	10

Table 6

ANOVA of Trials to Criterion on the Training Phase

For 19 Subjects:

Source	df	SS	MS	F	p
Total	56	4274.035			
Subjects	18	848.035			
Display	2	290.67	145.33	1.669	0.2027
Subjects x Disp.	36	3135.33	87.093		

For 26 Subjects:

Source	df	SS	MS	F	p
Total	77	6596.99			
Subjects	25	1208.32			
Display	2	598.79	299.39	3.125	0.0526
Subjects x Disp.	50	4789.87	95.797		

Mean Trials to Criterion

n	Pacer	Concrete HUD	Abstract HUD
$\bar{X}$(n=19)	15.63	16.32	20.89
$\bar{X}$(n=26)	15.08	18.346	21.85

Scheffe Test on 26 Subjects

Contrast	$d^{*}_{obtained}$	$d_{crit,.0526}$	p
Pacer - Concrete HUD	= -3.27	6.786	ns
Pacer - Abstract HUD	= -6.769	6.786	ns

$^{*}d = {}_{i} - {}_{j}$

Table 7

ANOVA of Trials to Criterion on Testing Phase

Source	df	SS	MS	F	p
Total	56	1352.84			
Subjects	18	351.51			
Displays	2	149.16	74.58	3.151	0.0548
Subjects x Disp.		36	852.175	23.67	

	Pacer	Concrete HUD	Abstract HUD
$\overline{X}$(n=19)	8.79	11.37	12.47

Scheffe Test

Contrast	$d_{obtained}$	$d_{crit,.0526}$	$d_{crit,.1}$	P
Pacer - Concrete HUD	= -2.58	3.95	3.52	.10
Pacer - Abstract HUD	= -3.68	3.95	3.52	.10
Pacer-.5(Concr. + Abs.)	= 3.13	3.95	3.52	.10

Table 8

ANOVA of Percent Correct Digit Sequences on First Trial (FIRST)

Source	df	SS	MS	*F*	*p*
Total	56	14925.29			
Subjects	18	7477.18			
Display	2	599.43	299.72	1.58	0.2208
Subjects x Disp.	36	6848.68	190.24		

	Pacer	Concrete HUD	Abstract HUD
$\bar{X}$(n=19)	50.40	48.66	42.76

Table 9

ANOVA of Percent Correct Digit Sequences on Last Trial (LAST)

Source	df	SS	MS	F	p
Total	56	12362.11			
Subjects	18	6616.95			
Display	2	1009.3	504.65	3.84	0.0309
Subjects x Disp.	36	4735.87	131.55		

	Pacer	Concrete HUD	Abstract HUD
$\bar{X}$(n=19)	56.94	48.37	47.67

Scheffe Test

Contrast	$d_{obtained}$	d_{crit}	p
Pacer - Concrete HUD	8.58	5.54	.05
Pacer - Abstract HUD	9.28	5.54	.05
Pacer - .5(Concrete + Abstract)	8.93	5.54	.05
.5(Pacer + Concrete) - Abstract	4.99	5.54	.05

APPENDIX B

Appendix B

Breakdown of Subjects that Completed the Experiment

Order of Displays

	First	Second	Third	n
Group 1	Pacer	Abstract HUD	Concrete HUD	5
Group 2	Abstract HUD	Concrete HUD	Pacer	3
Group 3	Concrete HUD	Pacer	Abstract HUD	4
Group 4	Pacer	Concrete HUD	Abstract HUD	3
Group 5	Concrete HUD	Abstract HUD	Pacer	2
Group 6	Abstract HUD	Pacer	Concrete HUD	2
		TOTAL		19

Number of subjects that had the:

Pacer 1st	8
Abstract HUD 1st	5
Concrete HUD 1st	6

APPENDIX C

SUBJECT QUESTIONNAIRE

Name or initials (optional): ____________________________

Academic Major: ____________________

Age: ___

Sex: ____F ____M

Left or Right handed (circle one)

The following questionnaire is provided in order to obtain some of your feelings concerning the experiment in which you have just participated.

A. Did you feel aware of the task objectives for:

	YES	NO	Comments:
the flying tasks?	___	__	
the auditory task?	___	__	

B. Did you feel that the following displays were legible and easy to read?

	YES	NO	Comments:
Pacer desktop	___	__	
Linear tape HUD	___	__	
Digital HUD	___	__	

Please answer the following questions by circling the number that most accurately describes how you feel about the particular question.

1. How difficult was each of the displays to learn?

	VERY EASY ... VERY DIFFICULT
Pacer desktop	1--2--3--4--5--6--7
Linear tape HUD	1--2--3--4--5--6--7
Digital HUD	1--2--3--4--5--6--7

2. How difficult do you feel each of the maneuvers was to learn using the following displays?

The Pacer desktop:

	VERY EASY ... VERY DIFFICULT
Climb	1--2--3--4--5--6--7
Cruise	1--2--3--4--5--6--7
Descent	1--2--3--4--5--6--7
Right Turn	1--2--3--4--5--6--7
Left Turn	1--2--3--4--5--6--7

The Linear tape HUD:

Climb	1--2--3--4--5--6--7
Cruise	1--2--3--4--5--6--7
Descent	1--2--3--4--5--6--7
Right Turn	1--2--3--4--5--6--7
Left Turn	1--2--3--4--5--6--7

The Digital HUD:	VERY EASY	VERY DIFFICULT
Climb	1--2--3--4--5--6--7	
Cruise	1--2--3--4--5--6--7	
Descent	1--2--3--4--5--6--7	
Right Turn	1--2--3--4--5--6--7	
Left Turn	1--2--3--4--5--6--7	

3. Without the auditory task and averaged over all maneuvers, how would you rate the three displays based on the amount of workload you felt in flying to the preset criteria?

	VERY LOW WORKLOAD	VERY HIGH WORKLOAD
Pacer desktop	1--2--3--4--5--6--7	
Linear tape HUD	1--2--3--4--5--6--7	
Digital HUD	1--2--3--4--5--6--7	

4. With respect to the auditory task, how would you rate the following in terms of the amount of workload you experienced?

	VERY LOW WORKLOAD	VERY HIGH WORKLOAD
Auditory task alone	1--2--3--4--5--6--7	
Auditory task and:		
Pacer desktop	1--2--3--4--5--6--7	

Linear tape HUD	1--2--3--4--5--6--7
Digital HUD	1--2--3--4--5--6--7

5. When performing the auditory and the flying tasks together, how many odd-even-odd sequences do you feel you missed while flying the following displays?

	NONE	MANY
Pacer desktop	1--2--3--4--5--6--7	
Linear tape HUD	1--2--3--4--5--6--7	
Digital HUD	1--2--3--4--5--6--7	

6. How often do you play video games?

at least once: ___ per day. ___ less than once a month.

___ per week.

___ per month.

7. How would you rate your interest in flying:

	NO INTEREST	GREAT INTEREST
before the study?	1--2--3--4--5--6--7	
after the study?	1--2--3--4--5--6--7	

8. How helpful was the experimenter during the study?

NOT HELPFUL	VERY HELPFUL
1--2--3--4--5--6--7	

9. How difficult do you feel it was to keep the following flight parameters within the required range?

Using the Pacer Desktop:

	VERY EASY	VERY DIFFICULT
Vertical velocity	1--2--3--4--5--6--7	

Bank angle	1--2--3--4--5--6--7
Airspeed	1--2--3--4--5--6--7
Altitutde	1--2--3--4--5--6--7
Heading	1--2--3--4--5--6--7

Using the Linear tape HUD?

	VERY EASY	VERY DIFFICULT
Vertical velocity	1--2--3--4--5--6--7	
Bank angle	1--2--3--4--5--6--7	
Airspeed	1--2--3--4--5--6--7	
Altitude	1--2--3--4--5--6--7	
Heading	1--2--3--4--5--6--7	

Using the Digital HUD:

	VERY EASY	VERY DIFFICULT
Vertical velocity	1--2--3--4--5--6--7	
Bank angle	1--2--3--4--5--6--7	
Airspeed	1--2--3--4--5--6--7	
Altitude	1--2--3--4--5--6--7	
Heading	1--2--3--4--5--6--7	

10. If you had to use one of the simulated head-up displays, which display would you prefer to use?

___ linear tape HUD ___ digital HUD

Why:

Please add any additional comments, suggestions, or feelings you may have concerning the tasks, equipment, or experimenter (use reverse side if necessary):

END

FILMED

10-83

DTIC